3.0 Data for the Screening Assessment

This section explains the process that was used to gather data for the contaminants identified and provides the data that were used in the ecological (Section 4.2) and human health (Section 5.2) screening assessments. Section 3.1 briefly describes the scope of the data-gathering process, and Section 3.2 discusses the approach. In Section 3.3, the data-gathering process and sources are described. Section 3.4 describes the analysis process used to select concentration input for the screening assessment models from the raw concentration data. Section 3.5 discusses the data that were selected and the use of substitute data for missing data.

For the screening assessment, we needed to find information (monitoring data) about the

- **26 contaminants (28 when accounting for various constituents of contaminants for which data were available) potentially in**
- **4 media (groundwater, sediment, seeps, surface water) plus external radiation at**
- **27 segments (areas) along the Columbia River, including the Hanford Reach and downstream to the first dam (McNary).**

3.1 Scope

The Columbia River has been the focus of environmental monitoring programs for five decades. The scope of the data task for this screening assessment was to compile data collected by the various monitoring programs for the contaminants of interest (that is, which contaminants should be evaluated in this screening assessment). Because the scope of the assessment is the current state of the river, January 1990 was selected as the earliest date for which data would be collected. Data after January 1990 reflect both current conditions and high-quality monitoring methods.

For the ecological and human health calculations for the screening assessment, data on contaminant concentrations were needed for the following media: groundwater, sediment, seeps, and surface water (Figure 3.1). In addition, external radiation data were needed for the human health assessment. Some of the data available but not complete enough for assessment purposes are contaminant concentrations in biota, cobalt-60 particles, drive point groundwater data for chromium, N Springs punch point water data, and pore water data for chromium. These data were used in limited calculations and model validation exercises.

3.2 Approach

All defining decisions for collecting and processing the data were made with CRCIA Team concurrence. All team decisions relating to the efforts of the data task are presented in Table 3.1.

The data needed to be

- **found, gathered, and identified according to the segment in which the data originated (see Section 3.3)**
- **selected for use (see Section 3.4) in the screening assessment**

For each of these steps, we consulted and reached consensus with the CRCIA Team on how best to proceed.

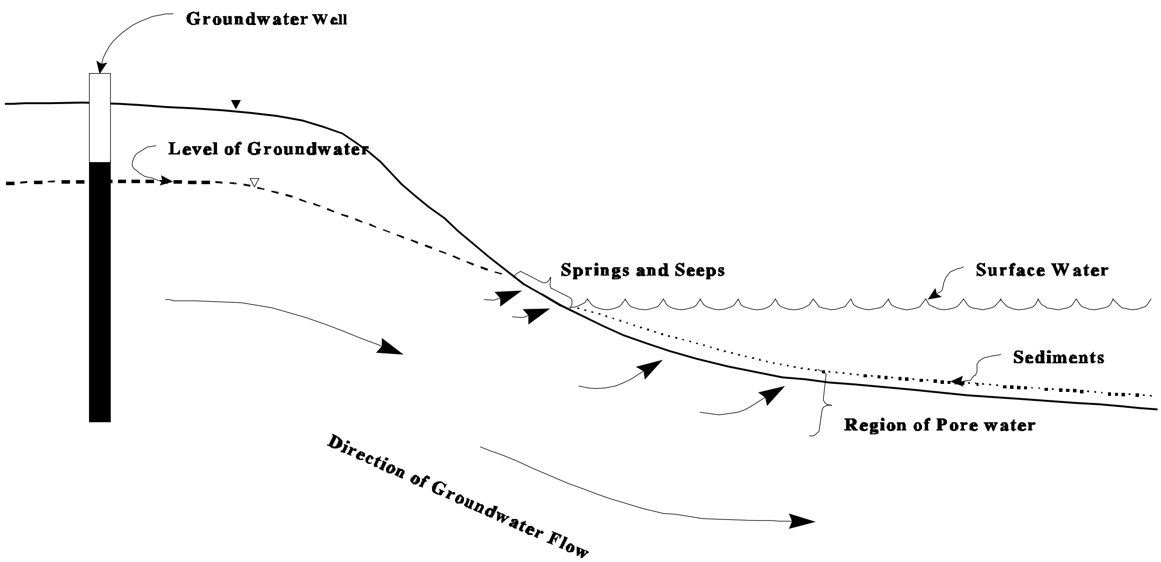


Figure 3.1. Groundwater, Sediment, Seeps, and Surface Water Media for Which Data Were Collected for the Screening Assessment

A Geographic Information System was used to help implement the processing of the data for the screening assessment. The Geographic Information System is a computerized system designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information. Many software packages perform basic Geographic Information System functions. Arc/Info Rev. 7.0.2 was used for the data task (ESRI 1994).

3.2.1 Segmentation

The ecological and human health screening assessments calculate potential risk based on contributions from multiple pathways affected by contaminant concentrations in multiple media. These contaminant concentrations were not usually measured in a fashion that would allow a complete assessment at every sampling site. To provide data for the assessments, data had to be aggregated to represent concentrations in areas rather than at points. This aggregation was done through river segmentation, a process that uses knowledge of contaminant sources and physical characteristics to localize risk results.

Staff from DOE, Ecology, and EPA defined 27 segments within the study area from Priest Rapids Dam down to McNary Dam (see Figure 3.2). A segment is a section of the river over which contaminant conditions can be expected to be similar and which captures the major influences to the Columbia River. The main resources used to decide how to most appropriately segment the river were as follows:

- a groundwater well location map (Dresel et al. 1995)

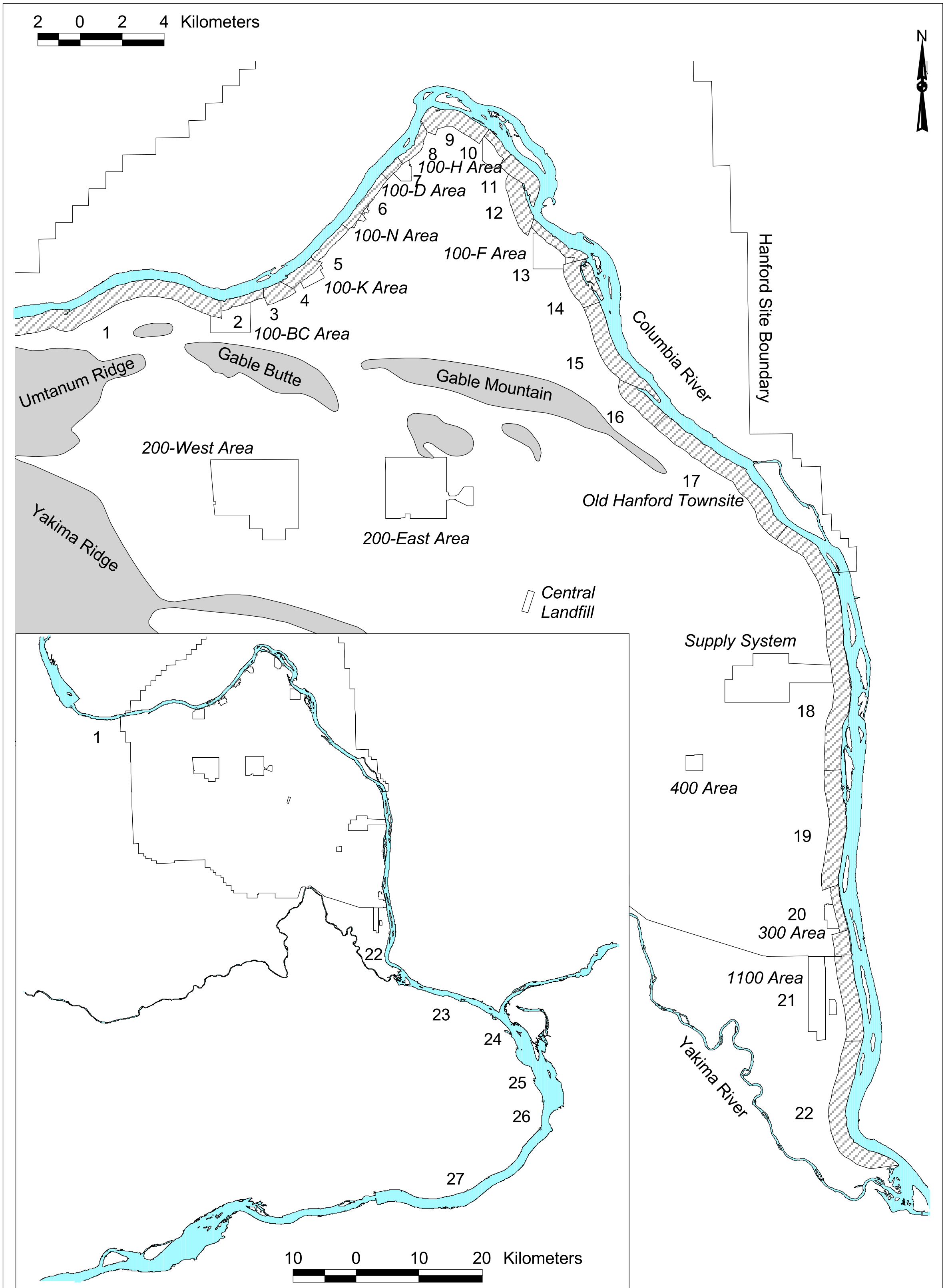
- the radiological and chemical contaminant plume maps from the 1994 Hanford Site groundwater monitoring report (Dresel et al. 1995)

Table 3.1. Data Decisions by the CRCIA Team

Date	Decision
1/30/96	Agreement was reached to collect data from January 1, 1990, to the present (June 1996) and to fill data gaps with older data where they were available for the initial phase of the screening assessment.
1/30/96	The primary geographic focus area for the screening assessment is from the vicinity of Priest Rapids Dam to McNary Dam. A rationale justifying this area will be provided by including in the report a discussion of historical levels/trends in contaminant data over time, showing levels typically upstream of McNary, including Hanford data, Oregon data, and Washington data.
2/13/96	All data will be provided on a diskette in the final report.
2/13/96	There will be no soil medium. Soil data will be generated from other media where needed. No soil samples are associated with the effluent pipe locations, and no other soil data were needed for the screening assessments.
2/13/96	The river (between Priest Rapids Dam and McNary Dam) will be broken into 27 segments. This partially defines the spatial aggregation of the data.
2/13/96	Corridor widths were chosen by segment based on sampling sites available to characterize contamination. Reactor areas 100 B/C, D, F, H, K, and N and the 300 Area have 0.4-km (1/4-mile) corridor widths. (N Reactor width was originally 0.8 km [½ mile] but changed to 0.4 km at 3/5/96 CRCIA meeting.) The non-trench portion of the 100-K Area has a 0.6-km (3/8-mile) corridor width. All other segments have a 0.8-km width. This completes the definition of the spatial aggregation of the data.
2/13/96	A representative value for each groundwater well in each segment will be chosen. A mechanized process needs to be developed to choose the representative value. The mechanized process is expected to be adequate for about 80 percent of the values. Remaining values will need to be examined by hand. A team was formed to develop the algorithm.
2/20/96	Where a clear upward or downward trend exists, a representative value will be chosen from the most recent data.
2/20/96	The maximum representative value for each data set should be an observed data point.
2/20/96	The set of representative data in each segment for each medium will be assumed to be lognormally distributed. The parameters for the lognormal distribution will be estimated from the representative data. Log-probability plots will be provided.
2/20/96	Data for both filtered and unfiltered water will be used in identifying representative data and in determining the parameters for the lognormal distribution.
2/27/96	Dixon's test will be used to eliminate, at most, a single outlier data point in each data set. In the data section of the final report, every data point that is eliminated will be explained.
2/27/96	To eliminate outliers, log transformation of the data will be used.
8/12/96	If no groundwater data are available, no other data will be substituted for the missing data. Substitute data would be used for sediment, seep water (groundwater data as a surrogate where available), and surface water. For surface water, if no measured data are available for Segment 1, Segment 2 data will be extrapolated if available. In Segments 2-27, data will be extrapolated from the nearest upstream segment with measured data.
10/1/96	The proposed system for substituting data when no sediment data are available is not workable, so no substitutions will be made.
5/14/96	Biota data, chromium drive point data from the Environmental Restoration Contractors (ERCs), and Westinghouse Hanford Company's (WHC) N-Springs punch point data were included for verifying the data used in the assessment.

File Contains Data for PostScript Printers Only

Figure 3.2. Segmentation of the Columbia River and Groundwater Corridors



segments.eps

- information from spring monitoring locations for the Surface Environmental Surveillance Project (unpublished information at PNNL)

- the maps of sampling locations for the special chromium studies being conducted by the Environmental Restoration Contractor at the 100-H Area (Hope and Peterson 1995) and 100-D Area (Hope and Peterson 1996)

- Ecology and EPA staff knowledge of the contaminant sources

Because many contaminant sources are located in the reactor areas, each reactor area was examined to determine if a single segment could be designed around it or whether it should be further divided. When this decision was made, the resources listed above were used to determine the upstream and downstream cutoff points for these segments. This process was continued for other major features of the river, such as the sloughs and the confluence points of the Yakima, Snake, and Walla Walla rivers. A slough is a backwater area that is along the shoreline and removed from the main body of the river. It is a unique environment unlike that of the neighboring river. In the segmentation process, each slough area was divided into two segments, one representing the river offshore from the slough and the other representing the slough itself. The distinction between the river and the slough segments is difficult to portray on a map (Figure 3.2). When creating the slough segments, each sampling location was reviewed by hand. The data owners and risk assessors (human health and ecological) were consulted to determine the most accurate placement of each sampling location in either the river or slough segment.

3.2.2 Thiessen Polygon Analysis

A Thiessen polygon analysis was used to monitor the appropriateness of the segment boundary placement. This was done by estimating the area of influence of a groundwater well and checking for two situations: 1) that all wells within a segment were estimating concentration for a similar size area, and 2) that boundaries reflected the area of influence of the wells contained within the segment. Thiessen analyses apportion points into polygonal regions such that each region contains only one point. Each region has the unique property that any location within a region is closer to the region's point than to any other point (Thiessen and Alter 1911). In this study, the points used for the analysis are the groundwater well locations.

Adjustments to segment boundaries were based on the Thiessen analysis for two reasons. First, it was desirable for the polygons within a segment to be of similar size so that each well would represent the contaminant concentration over an area of similar size. When one polygon became larger than the other polygons in the segment, the large polygon was clipped and the space was added to neighboring segments. This phenomenon occurs with the outer-most wells in a data-dense area when that data-dense area borders an area with few or no wells. The additional space was then added to a data-sparse segment. Figure 3.3 shows an example of this type of adjustment for the upstream boundary of Segment 18.

File Contains Data for PostScript Printers Only

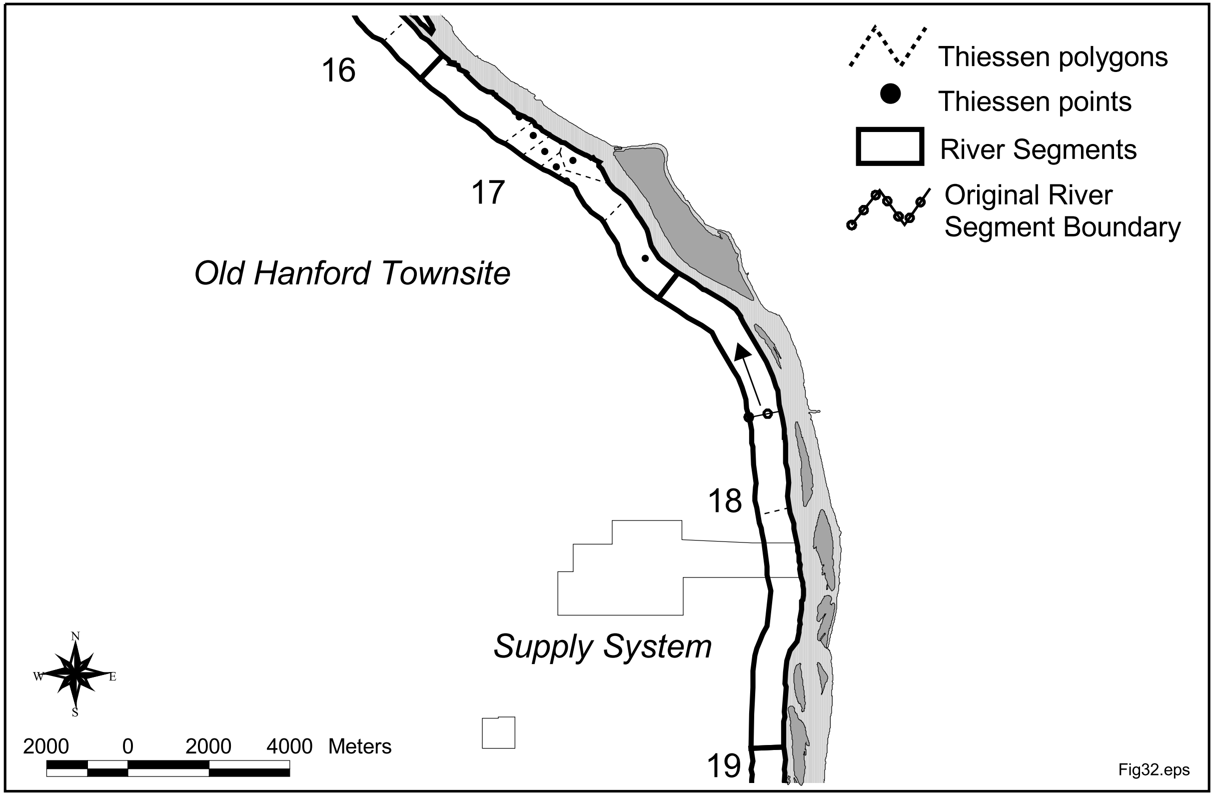
Figure 3.3. Example of Segment Boundary Adjustment Using the Thiessen Analysis to Reduce Data-Sparse Portions of Segments

Second, the original segment boundaries were drawn perpendicular to the river shore. To better reflect the areas of well influence, the polygon lines were used to represent the segment boundaries. Examples of this type of adjustment are shown in Figure 3.4 for both the upstream and downstream boundaries of Segment 3.

3.2.3 Corridor

A corridor was established in each segment to define the area from which data would be used for the various media. The corridor is defined by the length of the segment and a distance from the river shore. However, these corridors are not indicative of the spatial scope of the screening assessment calculations. The screening assessment applies to the Columbia River and the riparian zone.

Data for all media were initially gathered from a corridor up to 0.8 kilometer (½ mile) on either side of the Columbia River. For sediment, seeps, surface water, and external radiation, all data within 0.8 kilometer of the river were used. For the groundwater data, the number of wells close to the river was insufficient to monitor the groundwater concentrations entering the Columbia River. Therefore, a groundwater "corridor" was established for Segments 1-21 based on data available and knowledge of contamination sources. The corridor carried inland only to the point where a sufficient number of


Thiessen polygons
Thiessen points
River Segments
Original River
Segment Boundary
16
17
18
19
Old Hanford Townsite
Supply System
N
W
E
S
2000
0
2000
4000
Meters
Fig32.eps

File Contains Data for PostScript Printers Only

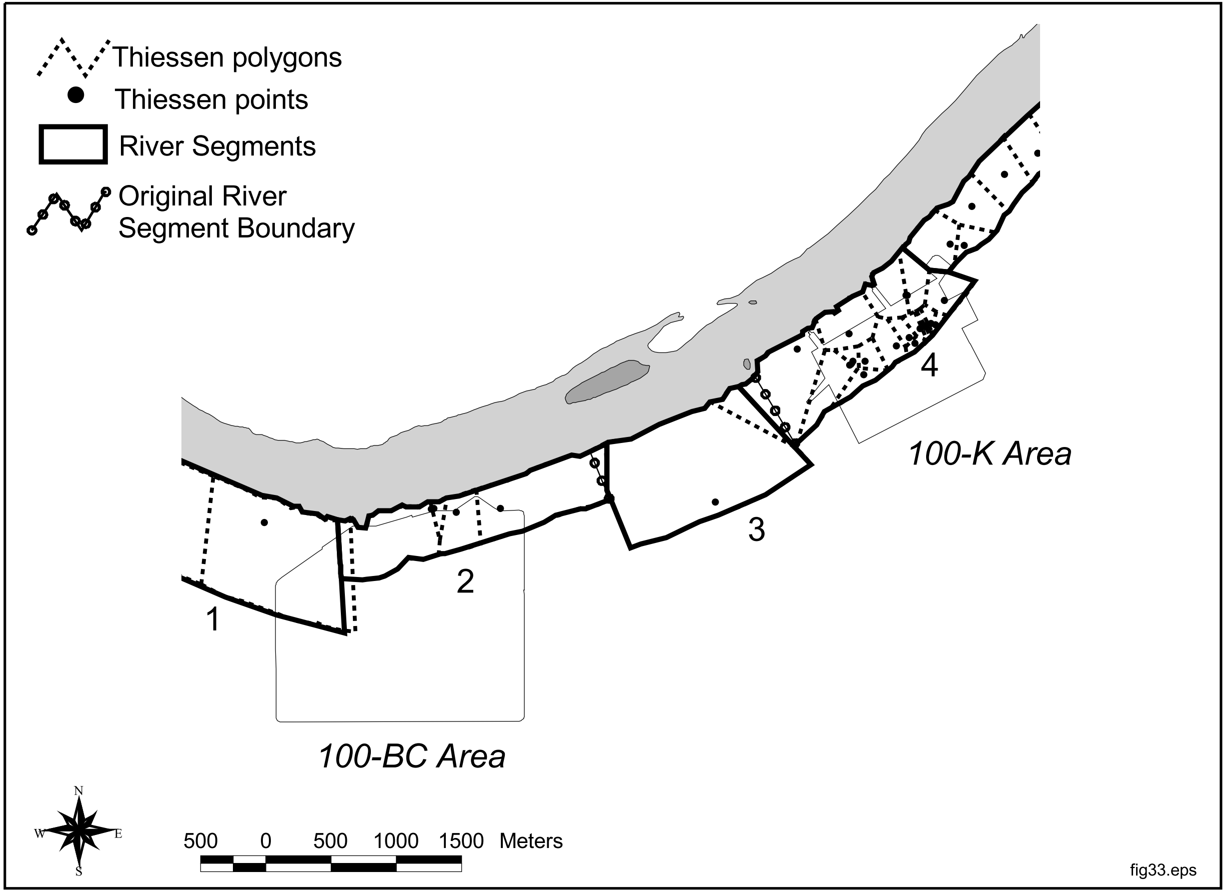
Figure 3.4. Example of Segment Boundary Adjustment Using the Thiessen Analysis to Better Reflect Areas of Influence from Wells

groundwater wells existed along the length of the segment to characterize contamination of the groundwater. No groundwater data below the City of Richland would be relevant to estimating Hanford impact on the Columbia River. The corridor width was based on having sufficient groundwater data to characterize the contamination within a segment. These corridor width decisions were made by staff from DOE, Ecology, and EPA with concurrence by the CRCIA Team. The corridor widths for groundwater data are as follows, as measured from the Columbia River shoreline inland onto the Hanford Site:

- Segments 2 (100-B/C Area), 5 (116-K-2 trench in 100-K Area), 6 (100-N Area), 7 and 8 (100-D Area), 10 (100-H Area), 13 (100-F Area), and 20 (300 Area) have a 0.4-kilometer (1/4-mile) corridor width.

- Segment 4 (100-K Area) has a 0.6-kilometer (3/8-mile) corridor width.

- All other segments have a 0.8-kilometer (½-mile) width (Figure 3.2).



Thiessen polygons
Thiessen points
River Segments
Original River
Segment Boundary
1
2
3
4
100-BC Area
100-K Area
N
W
E
S
500
0
500
1000
1500
Meters
fig33.eps

3.2.4 Distribution Assumption

All environmental concentration data used in the screening assessment were assumed to be lognormally distributed. Environmental data tend to be positive and are often highly skewed. The assumption of lognormality also forces sampling of the upper tail of the distribution. This assumption tends to represent the observed data well when a large number of concentration measurements with consistent analysis procedures are available. The assumption is not as accurate where analysis techniques vary, detection limits vary, and only a small number of analyses are available.

Some examples of frequency histograms of observed data are given in Figures 3.5 through 3.8. The first two examples are the concentration frequency of occurrence data for chromium in groundwater (Figure 3.5) and surface water (Figure 3.6) at the 100-N Area (Segment 6). The second two examples are the concentration frequencies of occurrence for cobalt-60 in groundwater at the 300 Area.

The example for chromium in the groundwater had 464 measurements, all analyzed with the same analytical technique. The chromium surface water example had only 23 measurements, and 4 of these had contaminated laboratory blanks associated with them. The measurements of 18 µg/liter and 19 µg/liter were all associated with the contaminated laboratory blanks. No modifications were made to the data in the situation described for chromium, as shown in Figures 3.5 and 3.6. Upper data were eliminated only if they were identified as outliers. There was no procedure for eliminating values with associated lab blank contamination.

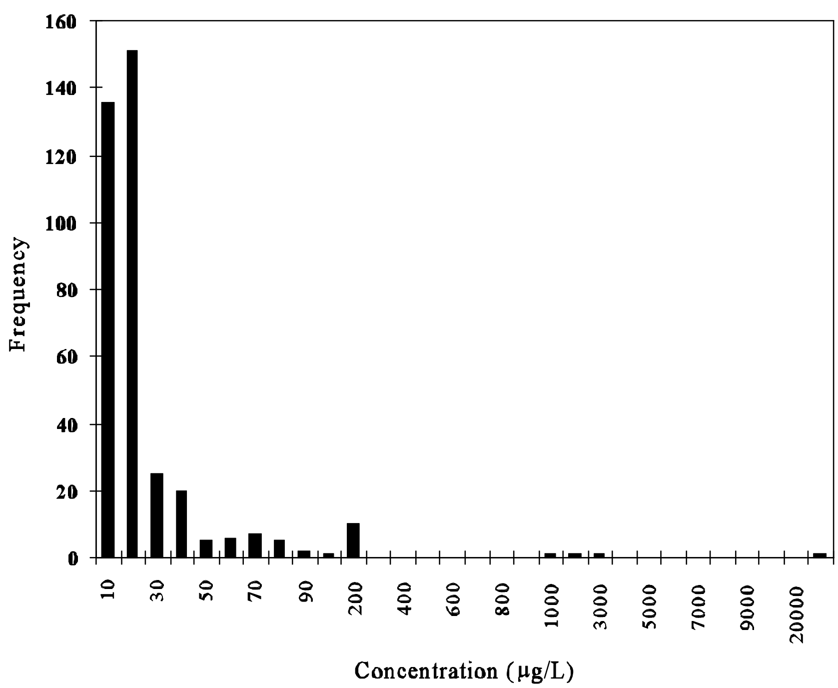


Figure 3.5. Frequency Histogram for Chromium in Groundwater at 100-N Area (Segment 6)

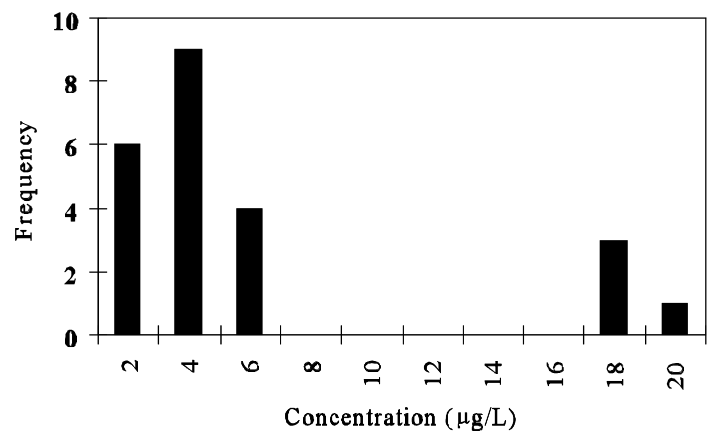


Figure 3.6. Frequency Histogram for Chromium in Surface Water at 100-N Area (Segment 6)

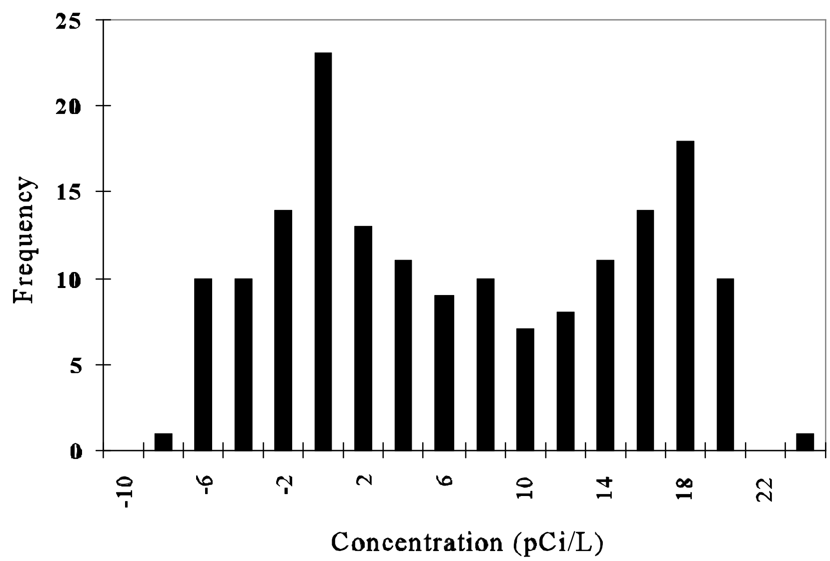


Figure 3.7. Frequency Histogram for Cobalt-60 in Groundwater at 300 Area (Segment 20) Depicting Two Detection Limits

The two examples showing frequency distribution for cobalt-60 in the groundwater at the 300 Area demonstrate the problems that occur when data are collected for different purposes. Figure 3.7 shows all of the observed data in the data set. The bimodal distribution of the data is due to a much higher detection limit for the data at the higher concentrations. These data were all flagged as undetected because the counting procedure was not designed to detect low concentrations of cobalt-60. Figure 3.8 shows the data set with only the data analyzed with the lower detection limit. For the radionuclide analyses, the data at the higher detection limit were removed from the data sets. This was done for all cobalt-60 and cesium-137 analyses where the analytical laboratory was given as “TMA.” Some concentrations for radionuclide measurements appear as negative numbers. For radiological analyses, the reported concentration is the measured concentration minus the concentration in a control source. When sample concentrations are near reference levels, the resulting reported concentration may be a negative number.

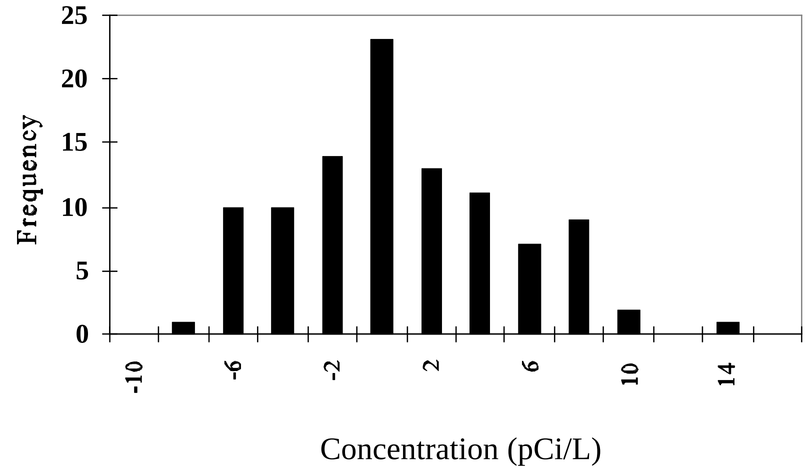


Figure 3.8. Frequency Histogram for Cobalt-60 in Groundwater at 300 Area (Segment 20) Depicting One Detection Limit

3.3 Data Gathering

Many data have been collected that are relevant to the Columbia River environment and CRCIA. The data task gathered the existing data that fell within the scope of the geographical area and time period of the screening assessment.

Both Hanford and non-Hanford organizations had data that met the data selection criteria of the screening assessment:

- **collected between the vicinity of Priest Rapids Dam and McNary Dam**
- **collected between January 1990 and August 1996**

Data were supplied by the City of Pasco, City of Richland, Hanford Site, Oregon State Department of Energy, U.S. Army Corps of Engineers, U.S. Geological Survey, Washington Public Power Supply System, Washington State Department of Ecology, and Washington State Department of Health.

One exception is the data for the effluent pipe system (see Section 2.3.2). Water used to cool the nuclear reactors was discharged into the central portion of the Columbia River via an effluent pipe system, which had pipelines buried in the river bottom. The liquid effluent passing through these pipelines has resulted in residual radioactive and chemical contamination. Several surveys of selected pipelines have documented chemically and radiologically contaminated scale on the inside of the pipelines and particulate contamination lying along the bottom of the pipelines. A video survey documented the structural integrity. Scrapings of the scale were analyzed for chemicals and radionuclides,